This session focuses on the estimation process for various types of models (except for logit models, which will be discussed in a later session). These models include cross-classification trip production models, regression models, gravity models, and aggregate time of day factor models. At the end of the series, participants will be able to better manage model development done for them by others (e.g., consultants) and to understand and evaluate the results.

Questions and Answers during Session III

Q: Why aren't the HBS trips made for dropping a child off included in the HB pickup category?
A: The driver would have a HB pickup/dropoff trip while the student would have a HB school trip.

Q: (From the homework) For JTW attractions, even if not employment, there would still be 560 attractions?
A: The constant is actually not statistically different from zero. It would be eliminated in application.

Q: (From the homework) Is the constant 560 too high?
A: We should not worry all that much about the intercept constant here because, as you will notice, the constant is insignificant (as shown by the t-stat).

Q: How is this coefficient calculated for nation in total?
A: The national coefficient is more an average across various urban areas, and not necessarily estimated. It can, however, be estimated using NHTS national sample data.

Q: What are variables not typically recommended? Like income etc
A: Usually variables that are extremely difficult to forecast.

Q: Is there a standard practice for applying weights based on trips/samples?
A: Typically weights are computed at the household level. This is because control totals are available at the household level. That is, data are available from national and local sources to tell us how many households of each type exist in the entire region. Once the household weights are derived, these can be applied to all persons in the household and for each trip made by a person in that household. Does this answer your question?

Q: NAS-TRB SR-288 identified fundamental shortcomings with aggregate/4-step/trip-based/ static assignment models. How can model estimation best enhance such modeling philosophy, as compared to advantages provided by tour-based models?
A: This is an excellent question. Unfortunately, the four-step models can never be fully enhanced to match the advantages provided by tour-based models. For instance, the issue of interconnectivity between trips and the transportation modes can never be addressed using a four -step model. The bottom line is that the MPO/urban area will need to determine the uses of the model, and determine if a four-step model will suffice or if a tour-based model is warranted.

Q: For the time of day where do you get those "simple factors" for the periods?
A: These are once again obtained from household surveys. Our Homework 3 deals with exactly this issue!

Q: How are weights determined?

A: Weighting is a complex statistical procedure. Typically, surveying firms obtain control totals for households based on several dimensions like geography, household lifestyle etc. Then, they use iterative proportional fitting procedures to develop weights.

Q: Due to the limitation of choosing only two variables in cross classification, why not use regression model?
A: Excellent point. We do use regression models for attractions because these are based on continuous variables like employment in a zone. Trip production uses cross-classification because the variables are household size, workers, vehicle etc. which are categorical variables. Having said that, there are advanced statistical models like Ordered Probit models that can model the choice of number of trips based on various household variables.

Q: What is an appropriate # of samples per cell?
A: Usually, we try to have at least 30-50 samples per cell. This is again a very subjective issue.

Q: Can we have cross classification models more than two dimensions?
A: Absolutely

Q: How would you classify volunteer work?
A: That probably will have to be home-based other trip, because it is discretionary and is not necessarily done on a regular basis.

Q: After you have estimated Productions and Attractions, and they come out different, which side do you better trust?
A: Very good question. Usually for home-based models, we tend to trust the production side more. For the NHB, we retain the production total, but use attractions to apportion the total into individual zones.

Q: For trip production modeling, a valid alternative structure is a logit model estimating the daily trip frequency a person (not a household) would make. The reasons why such model structure would be superior is because the logit structure allows more independent variables to be used; allows continuous independent variables to be used, rather than only classification variables; and allows statistical measures to be made measuring the significance of the independent variables and the entire equation. Will this be discussed at the next seminar?
A: Actually, the right way to estimate a discrete model for trip productions is an ordered probit model. Logit models can be used to estimate this, but a basic MNL will simply not be a good methodology, given the obvious correlations between various alternatives (levels of trip making). An ordered probit model is beyond the scope of this course, but if you are interested, I would recommend looking at the Discrete Choice Manual by Profs. Koppelman and Bhat.

Q: In case of Cross Classification method, there is no way to check the statistical significance of the variables other than logical check.
A: Absolutely right. It is a very simplistic model.

Q: Would it be a good approach to estimate at TAZ and fill underestimated zones by districts aggregation?
A: This would be a good approach if the TAZs that are included in the survey are indeed representative of travel behavior to and from that particular TAZ. However, this is hardly the case in household surveys. The reason behind aggregation is not only data limitation but also data representativeness. It is more likely that macroscopic patterns are captured well in a

survey. That's why, for attraction models in particular, aggregating to super districts may make sense.

Q: Is simple linear regression the only regression used for trip attraction?
A: Actually more advanced models like Attraction End choice models have been used both in theory and practice. Such models completely preclude the need for linear-regression based models for attractions.

Q: There is a concern in using trip productions for balancing. Into the future if households grow more than that of employment, then in the future we get more productions than attractions. Then the low growth of employment is not represented when the trips are balanced based on productions.
A: Good point. Indeed, balancing presents all sorts of problems, like the ones you pointed out. This is why, moving away from production and attraction framework to production and attraction end choice may be a good ides. In such a case, you would model the productions and then apportion these productions to attraction ends based on a solid discrete choice model that is based on underlying employment characteristics. There is no need to forcefully balance productions and attractions.

Q: Has ordered probit modeling been successfully implemented in any US MPO/COG's trip production models?
A: Not to my knowledge. I have seen models that are either simple cross-class based production models or tour-based models that model that don't need a trip production model. But going back to your original point about logit models being used for trip production, one can use a multiple nested logit model to represent various levels of trip making.

Q: Are the Attraction End Choice models logit models?
A: Yes. These are logit models with each choice being a zone.

Q: Is there guidance for minimum sample size for each trip purpose in the data for time of day factors?
A: Not really...Again, we make sure there are at least 20-30 records for each bin.

Q: Is auto ownership a good proxy for income in the cross classification model?
A: Excellent point. Yes. Indeed many cross-class models do use auto-ownership instead of income.

Q: Could you please repost the next class sessions?
A: Please check the calendar on the TMIP Web site at http://tmip.fhwa.dot.gov/.
A: February 10, 2009, Travel Modeling Workshop, Session 4; March 12, 2009, Travel Modeling Workshop, Session 5.

Q: For trip distribution, is the Census LEHD data a good source for adjusting the trip table?
A: Sure...Census data can be used to validate trip distribution models estimated from local survey data for HBW trips.

Q: Please explain logsums.
A: I can't do it very quickly, but it will be quicker in the next session. It requires some knowledge of logit models, which we will cover next time.

Q: How is OD survey used in validating friction factors?

A: OD surveys give us a way of calculating trip lengths, and matching our model estimates to these trip lengths.

Q: Do you think that auto ownership is a better predictor rather than income in the cross-class model?
A: Excellent point. Yes. Indeed many cross-class models do use auto-ownership instead of income.